Eur. J. Clin. Chem. Clin. Biochem.
Vol. 30, 1992, pp. 415–424

Rationale for Using Multiple Regression Analysis with Complex Interferences

By *M. H. Kroll* and *Ruth Chesler*

Clinical Pathology Department, Clinical Center, National Institutes of Health, Bethesda, Maryland, U.S.A.

(Received April 22, 1991/March 11, 1992)

Summary: Non-specificities and interferences may become complex when they involve the analyte as well as other interfering substances. These non-specificities and interferences are known as analyte-dependent and multi-interferent interferences. Multiple regression analysis has proven valuable in analysing this type of interference, but the theoretical foundation for using multiple regression analysis to study the basic mechanisms of interference has not been explicitly demonstrated.

Graph theory can depict and model the basic mechanisms of interferences and the possible interactions. The relationship between the analyte, the interferents, and the response of the instrument to these entities can be approximated by a polymial of order three, which includes partial derivatives and cross-terms. The partial derivatives relate to the different interactions found with the graph theory model. Further, the partial derivatives can be associated with the coefficients in the multiple regression analysis when the respective values of the three variables (analyte, interferent one, and interferent two) are multiplied by one another. One can decide to retain or discard the coefficient of a variable, based on the statistical significance of the coefficient. The respective interactions in the graphic model can then be assembled and the framework of the interference mechanism established.

Introduction

Non-specificities and interferences may become complex when they involve the analyte as well as other interfering substances. These non-specificities and interferences are known as analyte-dependent and multi-interferent interferences. Multiple regression analysis has proven valuable in analysing this type of interference. Failure to use multiple regression analysis may lead one to the false conclusion that the interference is independent of the analyte. Interferences dependent on the analyte have been demonstrated (1). In addition, two different interferents may interact with each other, independent of the analyte, and result in multi-interferent interference, as shown by the negative-interference in the determination of total protein or albumin caused by the combined effects of bilirubin and salicylate (2). Finally, the analyte and two interferents may interact, resulting in a complex interference. Only proper analysis can fully characterize the presence or absence of interactions between analyte and interferent or multiple interferents.

Multiple regression analysis provides considerable information concerning the interaction between analyte and interferents or multiple interferents using straightforward manipulations of analyte and interferent concentrations. Procedures for the application of regression analysis for analyte-dependent and multi-interferent interference models have been previously recommended (1, 3). Previous studies presented details of the experimental setup and interpretation of the results (1), but none of these studies rigorously demonstrated the rationale of using multiple regression analysis with analyte-dependent or multi-dependent interference, a simple procedure for handling the data, or the theoretical foundation connecting the basic

mechanisms of interference with the multiple regression model. Here we present an overall scheme of the basis interactions of interferences with the analyte, using graph theory. We express the response of the system as a function of the concentrations of analyte and interferents and the partial derivatives of the transformation function with respect to the components (concentrations of analyte and interferents). The partial derivatives are equated with the connections of the graphic model and the coefficients of the multiple regression analysis. Further, we clarify the nomenclature and manner of data entry.

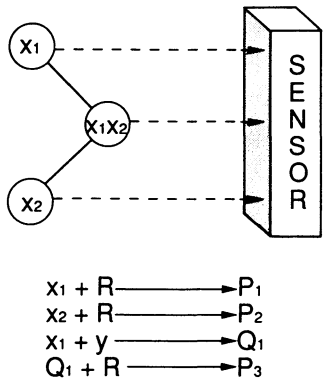

Fig. 1. Graphic model of the interactions between analyte (x_1) and one interferent (x_2). Both the x_1 and x_2 can interact independently with the sensor, and their interaction complex ($x_1 x_2$) interacts with the sensor too. The reactions are shown in the lower half without regard to stoichiometry, with R representing the reagents, P representing an interaction with the sensor, and Q representing the product between the analyte and the interferent.

Graph Theory Approach

One purpose of studying interferences is to determine the magnitude of the interference in terms of the interferent concentration. It is possible to estimate the significance of the interference in a given sample and at a known concentration of interferent. Another purpose is to determine the mechanism of the interference. Knowledge of the mechanism of the interference allows one to decrease the effect of the interferent by modifying the method. Thus, elucidation of the mechanism, although difficult, is critical for optimizing new methods (4). Multiple regression studies can help to elucidate the mechanism by determining the general class of interference.

In general, interference mechanisms are analyte-dependent, analyte-independent, or multi-interferent. Graph theory is a tool for solving combinatorial problems and can thus elicit the various possibilities for the combination of the different chemical species of interest (5). Graph theory represents the chemical species as nodes (circles with letters in them) and the possible relationships as edges or connections (the connecting lines) (fig. 1). Figure 1 illustrates the simplest case, with x_1 as the analyte, x_2 as the interferent, and S as the sensor or detector; one can construct the possible combinations for analyte-dependent and analyte-independent interference. We assume that the reagent, R, has a high enough concentration to react with the analyte, the interferent, or the analyte-interferent combination without significant change in concentration. The reagent can be as simple as water in a colorimetric procedure, such as a bilirubinometer, or the flame of flame photometry and atomic absorption. Of course, the reagent can be much more complicated and the analytical reaction can invoke many steps, such as those involved in enzyme-linked reactions. In reactions with many steps, the analyte, x_1, may represent one of the intermediates or products of this chain of reactions. The interferent, x_2, may react with any of the products in the chain yielding the analyte-dependent species, $x_1 x_2$. Thus, in this model the interferent is not restricted to reacting only with the analyte.

A reagent giving rise to the product, P, interacts with the sensor, with either positive or negative effects. Figure 1 shows this interaction as a dashed line with the arrow pointing to the sensor. All the possible interactions with one interferent are shown. As always, the analyte reacts with the reagent and then interacts with the sensor to produce a response. The response is then translated into a number meaningful to the observer. The interferent can also react with the reagent and then interact with the sensor. The analyte and the interferent may interact forming a new compound or complex, Q_1. This new compound or complex, Q_1, may interact with the reagent and form a new product, P_3, which interacts with the sensor creating a signal. In this way, a positive analyte-dependent interference occurs. If Q_1 does not form P_3, then a negative analyte-dependent interference ensues. Almost all analytical reactions in clinical chemistry are kinetic or endpoint (equilibrium) or some variant of these two. For a kinetic method, the equation stipulating the reaction can be written as $dQ_1/dt = k x_1 x_2$, that is the rate of change equals a rate constant multiplied by the analyte concentration multiplied by the interferent concentration. For an equilibrium reaction, $K = Q_1/x_1 x_2$, where K is the equilibrium constant. Thus $Q_1 = K x_1 x_2$. Thereby, the designation of an interaction as $x_1 x_2$ within a circle is justified.

The simple case of two interferents differs from that of a single interferent. Here the analyte (x_1) does not interfere with either interferent (fig. 2). But the two interferents, x_2 and x_3, respectively, do interact and form a new chemical, $x_2 x_3$, designated as Q_2 in figure 2. Each of the interferents may react with reagents and interact with the sensor. The new chemical, Q_2, may also react with the reagents and interact with sensor, or it may simply remove the interferents from the reaction medium.

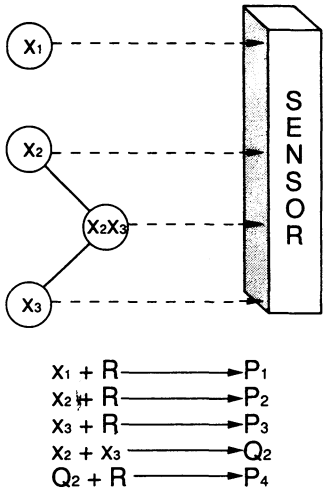

Fig. 2. Graphic model of the interactions between one interferent (x_2) and a second interferent (x_3). The analyte (x_1) does not interact with the interferents. The interferents, x_2 and x_3, can interact independently with the sensor, as does the analyte, x_1. In addition, the complex of interferents one and two, ($x_2 x_3$), can interact with the sensor. The reactions are shown in the lower half without regard to stoichiometry, with R representing the reagents, P representing an interaction with the sensor, and Q representing the product between the interferent-one and the interferent-two.

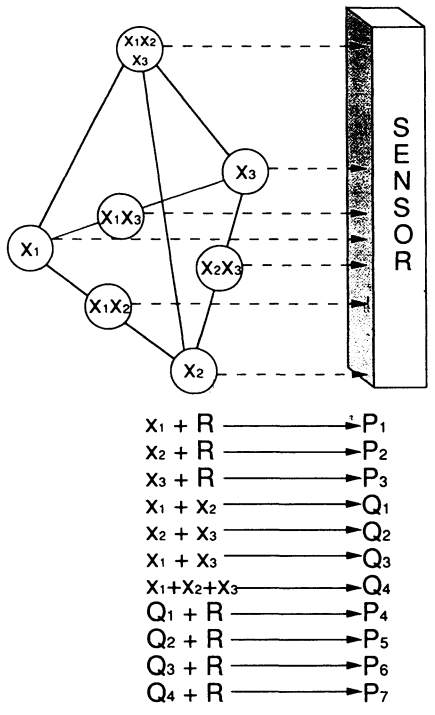

Fig. 3. Graphic model of interactions between and among analyte (x_1), interferent-one (x_2), and interferent-two (x_3). The graphic diagram forms a pyramid of possible interactions. Analyte (x_1) and interferents (x_2 and x_3) can interact directly with the sensor. They form the base of the pyramid. Analyte and interferents can pair with one another to form complexes ($x_1 x_2$, $x_2 x_3$, and $x_1 x_3$), which in turn have the potential to interact with the sensor. The analyte and interferents can form a three-way complex ($x_1 x_2 x_3$) that has the potential to interarct with the sensor. The reactions are shown in the lower half without regard to stoichiometry, with R representing the reagents, P representing an interaction with the sensor, and Q representing the products between the analyte and the interferents.

The basic reactions, without constraints on the reactivity of the analyte, x_1, for the presence of two interferents, are more complex than presented above. Here the analyte may react with both interferents, x_2 and x_3, to give the respective products, $x_1 x_2$ (Q_1) and $x_1 x_3$ (Q_3). In addition, the two interferents may react with each other to form the product $x_2 x_3$ (Q_2) as shown in figure 3. In addition, the analyte may react with both interferents to form the product $x_1 x_2 x_3$ (Q_4). This reaction appears to be trimolecular, which is highly unlikely; however, one must recall that the sensor detects changes over the entire time-course of the reaction and that the kinetics of the reaction may be very quick. The product $x_1 x_2 x_3$ may result from an initial reaction of x_1 with x_2 to form $x_1 x_2$, followed by the reaction of $x_1 x_2$ with x_3 to form $x_1 x_2 x_3$. Similarly, the product may occur if x_1 initially reacts with x_3 to form $x_1 x_3$, followed by the reaction with x_2 to form $x_1 x_2 x_3$, or if x_2 initially reacts with x_3 to form $x_2 x_3$ followed by the reaction with x_1 to form $x_1 x_2 x_3$. One would expect such a reaction to occur if one of the chemical species is an enzyme or macromolecule, such as a protein, lipid, or nucleic acid. Any and all of the four newly formed products may react with the reagents and interact with the sensor. Failure to interact with the sensor results in a negative interference.

The cases examined so far have included only those where each of the three chemical species, x_1, x_2, and x_3, reacts only with components different from itself. The situation where species react with themselves can be called autoreactivity and pictured as x_1 with x_1 to from $x_1 x_1$ (A_1), x_2 with x_2 to form $x_2 x_2$ (A_2), and x_3 with x_3 to form $x_3 x_3$ (A_3) (fig. 4). These products could go on to react with yet another species to form a trimolecular product, $x_2 x_2$ with x_1 to form $x_1 x_2 x_2$ ($C_{1,2}$, where the subscripts indicate that one x_1 component and two x_2 components are present in the complex) (fig. 4). Autoreactive reactions are rare in the analytical methods employed in clinical chemistry.

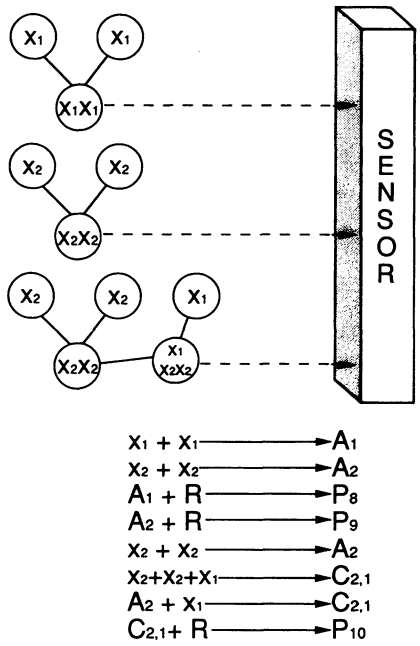

Fig. 4. Graphic model of autoreactivity among analyte (x_1) and interferent (x_2). Analyte may react with itself to form an autoreactive complex ($x_1 x_2$). Interferent may react with itself to form an autoreactive complex ($x_2 x_2$). The autoreactive complex of the interferent ($x_2 x_2$) may interact with the analyte to form an analyte-interferent-interferent complex ($x_1 x_2 x_2$). All of these complexes have the potential to interact with the sensor. The reactions are shown in the lower half without regard to stoichiometry, with R representing the reagents, P representing an interaction with the sensor, A representing autoreactive complexes of either the analyte or the interferent, and C representing the complex formed from the interferent autoreactive complex and the analyte. The C complex can take more than one pathway. Even though autoreactive complexes are possibly, they have rarely been of significance in the interferences studies thus far.

In general, these analyte and interferents may be thought of as elements of the system. Each element may react with itself (an autoreactive or autocatalytic reaction), or with the other elements. The order of interaction represents the number of elements interacting. The possible different combinations, which represent the global mechanism of mechanisms, at each order of interaction, is given by the permutation of these given elements (5). The permutations, and thus the possible interactions, can be determined using graph theory.

Two Examples

These ideas become clearer when applied to examples. The first example originates from the NCCLS guidelines for interference (3). The two fractions of bilirubin, conjugated bilirubin and unconjugated bilirubin,

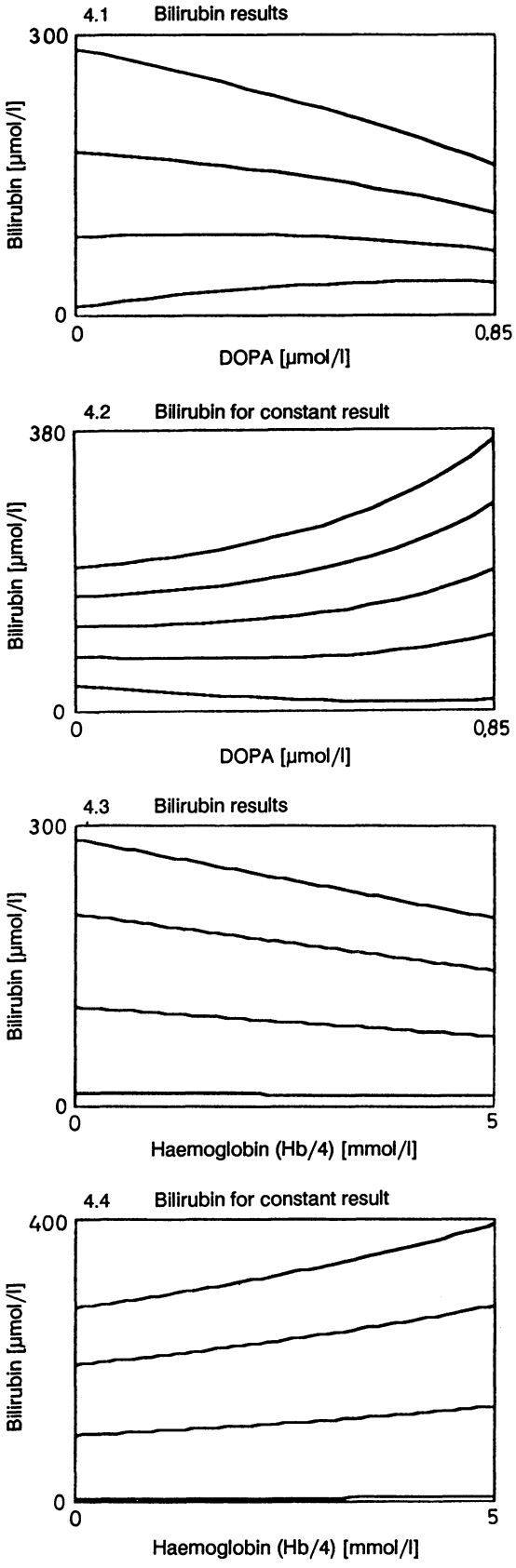

Fig. 5. *Contour* plots of the interference of DOPA or haemoglobin with bilirubin. The first plot shows the bilirubin results as a function of DOPA for varying bilirubin concentrations. The second plot shows the bilirubin concentration necessary to obtain a constant bilirubin result as a function of DOPA at varying bilirubin concentrations. The third plot shows the bilirubin results as a function of haemoglobin for varying bilirubin concentrations. The fourth plot shows the bilirubin concentration necessary to obtain a constant bilirubin result as a function of haemoglobin at varying bilirubin concentrations.

both interfere with the method for γ-glutamyltransferase. The interaction between conjugated and unconjugated bilirubin also appears to interfere, as shown by the data in table 2. The interaction of conjugated and unconjugated bilirubin may be in the form of a single molecular complex, or each of them may affect the reactions leading to the sensor at different steps along the pathway (fig. 2).

For the second example, we examine the effect of *L*-DOPA and haemoglobin (Hb) on total bilirubin (bilirubin) as measured with the Kodak Ektachem (Eastman Kodak Co., 225 East Ave., Rochester, NY 14604, USA). Bilirubin as bilirubin reference material and human haemoglobin (crystallized, dialysed and lyophilized) were obtained from Sigma Chemical Co. (P. O. Box 14508, St. Louis, MO 63178 USA). DOPA was obtained from Aldrich Chemical Co. (1001 West Saint Paul Ave., Milwaukee, WI 53233 USA). We varied independently the concentrations of bilirubin, DOPA, and haemoglobin, ensuring that there were four different concentrations for each (zero for the baseline of DOPA and haemoglobin).

Analysing the interaction as outlined in the previous section and in figure 3 presents us with the possible interactions for the bilirubin and the two interferents. But the problem becomes exceedingly complicated without an a priori knowledge of the mechanisms of interference. We therefore need an experimental approach that will guide us to the probable mechanisms; response surface modelling using multiple regression analysis offers such an approach.

Response Surface Model and the *Taylor* Expansion

Another approach to model building besides the combinational one is to assume that one knows nothing about the underlying mechanism of the interactions, but that there exists a functional relationship between the different variables in the system, i. e. the analyte and interferent concentrations and the ultimate output of the instrument (6). Such a functional relationship can be expressed as $\eta = g(\xi_1, \xi_2, \ldots, \xi_k)$, where ξ is the vector of variables (6). Because one cannot know this function exactly, one must approximate it by an empirical polynomial including all combinations, as shown below for a two-variable system:

$$\begin{aligned} g(\xi) = g(\xi_1, \xi_2, \ldots, \xi_k) &= f(\mathbf{x}, \beta) \\ &= + \beta_0 \\ &\quad + (\beta_1 x_1 + \beta_2 x_2) \\ &\quad + (\beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2) \\ &\quad + (\beta_{111} x_1^3 + \beta_{222} x_2^3 + B_{112} x_1^2 x_2 + \beta_{122} x_1 x_2^2) \\ &\quad + \text{etc.} \end{aligned}$$

(6). After analysing the results by multiple regression (using the method of least squares), one can relate results (output) to the variable concentrations (inputs) (6). When a set of variables is written as $x_1 x_2$, that expression represents an interaction term. If the analysis warrants that $x_1 x_2$ be kept in the full expression, this means that the combined interaction of x_1 and x_2 has an effect on the output and that this effect is independent of the effects of either x_1 or x_2 alone. When a set of variables is written as $x_1 x_1$ or x_1^2, that expression represents an autoreactive or autocatalytic term. If the analysis warrants that x_1^2 be kept in the full expression, it means that the combined interaction of x_1 with itself has an effect on the output and that this effect is independent of the effects of x_1 by itself. In other words, $x_1 + x_1 \rightarrow x_1^2$, and x_1^2 has an effect on the system separate and different from x_1 by itself.

In calibrating an instrument, one sets up a transformation that translates the electrical signals of the sensor to a concentration value of the analyte (7). This transformation, in it simplest form, can be expressed as the partial derivative with respect to the analyte, $\frac{\partial f}{\partial x_1}$ (8). *Pszonicki* extended these concepts to deal with the problems of linear and change-in-slope calibration curves (9, 10). Such an approach includes the interferents and non-specificities in the calibration step.

Once can extend the concept to include the effects of interferents on the mathematical transformation. The *Taylor* expansion of this true function with respect to the pertinent variables gives rise to the polynomial that best approximates the response-surface relationship. The response of the system can be thought of as being represented in the function $f = f(\mathbf{x})$, where $\mathbf{x}$ is the vector $(x_1, x_2, \ldots, x_m)^T$, the T indicating the transpose of the vector $\mathbf{x}$.

If one assumes that the response between the instrument and analyte and interferent concentrations is continuous, and that the transformation that describes this relationship is differentiable to degree n, then the *Taylor* expansion is given by

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}_0) &+ [(\mathbf{x} - \mathbf{y}_0) \cdot \nabla] f(\mathbf{y})|_{y = y_0} \\ &+ \frac{1}{2!} [(\mathbf{x} - \mathbf{y}_0) \cdot \nabla]^2 f(\mathbf{y})|_{y = y_0} + \cdots \\ &+ \frac{1}{(n-1)!} [(\mathbf{x} - \mathbf{y}_0) \cdot \nabla]^{(n-1)} f(\mathbf{y})|_{y = y_0} \\ &+ \frac{1}{n!} [(\mathbf{x} - \mathbf{y}_0) \cdot \nabla]^n f(\mathbf{y})|_{y = x_0} \end{aligned}$$

where **x** and **y** are vectors of the analyte and interferent concentrations and ∇ is the nabla operator $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_m}\right)$ (11). For the sake of simplicity, one can take $\mathbf{y}_0 = (\mathbf{0})$, thus for one analyte ($x_1$) and one interferent ($x_2$)

$$f(x_1, x_2) = f(0, 0) + x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_1 x_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + \frac{1}{2} x_1 \frac{\partial^2 f}{\partial x_1^2} + \frac{1}{2} x_2 \frac{\partial^2 f}{\partial x_2^2} + \text{H.O.T.}$$

The abbreviation H.O.T. stands for higher order terms of the continued *Taylor* expansion of the function f and the remainder. These terms are of the third and higher powers, $x_1^4 \frac{\partial^4 f}{\partial x_1^4}$ for example, and are more complicated than the lower powers. Including them may improve the accuracy of the expansion, but usually their values are so small in comparison with the lower power terms that they are excluded without loss of accuracy (12).

Comparing the truncated *Taylor* expansion with the response-surface expression

$$f(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2,$$

and matching like variables and orders, it is evident that

$$\beta_0 = f(0, 0),$$

$$\beta_1 = \frac{\partial f}{\partial x_1}, \beta_2 = \frac{\partial f}{\partial x_2}, \beta_{12} = \frac{\partial^2 f}{\partial x_1 \partial x_2},$$

$$\beta_{11} = \frac{\partial^2 f}{\partial x_1^2}, \quad \text{and} \quad \beta_{22} = \frac{\partial^2 f}{\partial x_2^2}.$$

Furthermore the relationships that were previously established between the combinational diagrams (fig. 1–4) and the response surface approach continue to apply; thus the partial derivatives relate to the appropriate β coefficient of each element and to the response of the system as mediated through the sensor. Thus, for the interferent x_2, $\beta_2 = \frac{\partial f}{\partial x_2}$, is the response of the instrument to x_2 while all other factors are held constant. Similarly for the cross reaction, $x_1 x_2$, $\beta_2 = \frac{\partial^2 f}{\partial x_1 \partial x_2}$ is the response of the instrument to the cross-reaction product of x_1 and x_2 while all other factors are held constant. Because the coefficients of the polynomial can be found by multiple regression analysis, the coefficients of the regression may be interpreted as the slopes of a particular element or particular interaction (tab. 1).

Tab. 1. Definitions of constants and their relation to terms in equations

Derivative	Coefficient	Associated variables
$f(0, 0, 0)$	β_0	none
$\frac{\partial f}{\partial x_1}$	β_1	x_1
$\frac{\partial f}{\partial x_2}$	β_2	x_2
$\frac{\partial f}{\partial x_3}$	β_3	x_3
$\frac{\partial^2 f}{\partial x_1 x_2}$	β_{12}	$x_1 x_2$
$\frac{\partial^2 f}{\partial x_1 x_3}$	β_{13}	$x_1 x_3$
$\frac{\partial^2 f}{\partial x_2 x_3}$	β_{23}	$x_2 x_3$
$\frac{\partial^3 f}{\partial x_1 x_2 x_3}$	β_{123}	$x_1 x_2 x_3$

The expansion for three variables is more complex. When one expands the equation for f to the second order and includes the third-order interaction term for all three variables, the expansion takes the form of

$$f(x_1, x_2, x_2) = f(0, 0, 0) + x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} + x_1 x_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + x_1 x_3 \frac{\partial^2 f}{\partial x_1 \partial x_3} + x_2 x_3 \frac{\partial^2 f}{\partial x_2 \partial x_3} + \frac{1}{2} x_1^2 \frac{\partial^2 f}{\partial x_1^2} + \frac{1}{2} x_2^2 \frac{\partial^2 f}{\partial x_2^2} + \frac{1}{2} x_3^2 \frac{\partial^2 f}{\partial x_3^2} + x_1 x_2 x_3 \frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3}$$

If two interferents are present but do not interact with the analyte, then one can simplify the analytical expansion of the response, by setting x_1 equal to a constant c; the function taking the form

$$f(x_1 = c, x_2, x_2) = f(c, 0, 0) + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} + x_2 x_3 \frac{\partial^2 f}{\partial x_2 \partial x_3} + \frac{1}{2} x_2^2 \frac{\partial^2 f}{\partial x_2^2} + \frac{1}{2} x_3^2 \frac{\partial^2 f}{\partial x_3^2},$$

because

$$\frac{\partial f}{\partial x_1}, \frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}, \frac{\partial^2 f}{\partial x_1 \partial x_3},$$

and

$$\frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3}$$

all equal to zero.

The first term on the right-hand-side of the equation, f(c, 0, 0), indicates that one has set the x_1 variable equal to the constant concentration c.

Application to Statistical Analysis

In regression analysis, values of coefficients are proposed that when multiplied with their respective variables minimize the error between the analytical values and the true values. The coefficients are calculated by solving the equation $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$, where **x** represents the vector of the intercept and the variables, **A** represents the matrix of values for the independent variables, the superscripted T represents the transpose, the −1 represents the inverse of a matrix, and **b** represents the experimental values of the dependent variable being fitted (13).

One must vary the concentration of the analyte, x_1, and the potential interferents, x_2 and x_3, so that they are linearly independent, and determine the value of the analyte with the appropriate instrument. After one has obtained the results, one should place the known values for each variable (analyte, interferent one, etc.) as well as the results of the determinations into its own column. Each cross-product term is considered a separate variable and has its own column; thus, one calculates the values for the cross-product term $x_1 x_2$ column by multiplying the x_1 by the x_2 values from the same row. We illustrate the set-up for the values for the NCCLS study in table 2: conjugated bilirubin (Bc), unconjugated bilirubin (Bu), Bu · Bc, and γ-glutamyltransferase each have their own column. We calculated the values in the Bu · Bc column by multiplying the respective values of conjugated bilirubin and unconjugated bilirubin in each row. We determined bilirubin in 64 samples in the bilirubin-DOPA-haemoglobin study (results for five samples are shown in the lower half of tab. 2), calculating the cross terms, Bilirubin-DOPA, Bilirubin-Hb, DOPA-Hb, and Bilirubin-DOPA-Hb, from the respective variables, Bili, DOPA, and Hb, for each row.

Tab. 2. Data set-up for multiple regression analysis

γ-Glutamyltransferase-conjugated bilirubin-unconjugated bilirubin

Tube number	Conjugated bilirubin (mg/dl)	Unconjugated bilirubin (mg/dl)	Bu · Bc	γ-Glutamyl-transferase (U/l)
1	0	0	0	100
2	10	0	0	90
3	20	0	0	80
4	0	10	0	95
5	10	10	100	85
6	20	10	200	70
7	0	20	0	90
8	10	20	200	75
9	20	20	60	60

A set value of 100 was added to the results from the NCCLS document. Bc = conjugated bilirubin, Bu = unconjugated bilirubin.

Abbreviated set-up for bilirubin-DOPA-haemoglobin interference

Bili	DOPA	Hb	Bili-DOPA	Bili-Hb	DOPA-Hb	Bili-DOPA-Hb	Results
4.3	0.43	0.53	1.8	2.7	0.23	0.98	27.4
80.4	0.22	1.6	17.3	129	0.35	27.8	82.1
170.2	0	1.6	0	272	0	0	164.2
170.2	0.86	1.07	147	182	0.92	157	107.7
282.2	0.86	0	243	0	0	0	145.4

Bili stands for bilirubin (μmol/l), DOPA for *L*-DOPA (mmol/l), and Hb for haemoglobin (g/l).

The multiple regression analysis provides us with several useful pieces of information. As shown in table 3, the correlation coefficient indicates the quality of the regression; if the correlation coefficient is low, there may have been a data-entry error or omission of an important variable. The F-test also indicates the quality of the regression. For the γ-glutamyltransferase-Bc-Bu interference study, the results are simple, and the coefficient and t-value are presented for each variable in table 3; from the t-value and the number of degrees of freedom one can determine the probability that the coefficient is not significantly different from zero. We recommend using a 5% level of significance ($p = 0.05$) for this test. Thus, any value of t less than 2.571 would not be significant. The value for the partial F test provides a criterion for discarding or retaining a variable in the model. Its degrees of freedom are one plus the degrees of freedom for the residuals. At the 5% significance level, a value for partial F greater than 6.61 would be considered significant. Thus, all three variables are significantly different from zero and contribute to the regression. The function that describes the interference is

$$f([Bc], [Bu]) = - [Bc] - 0.5[Bu] - 0.025[Bc][Bu],$$

where the brackets indicate the concentrations of these species. With regard to the mechanism, both conjugated bilirubin (Bc) and unconjugated bilirubin (Bu) decrease the reaction due to γ-glutamyltransferase, and when both are present a further reduction is encountered. This interference with γ-glutamyltransferase is an example of how two independent interferents can interact together and affect the analytical reaction.

The interaction of DOPA and haemoglobin in the determination of bilirubin using the Ektachem is more complex. We show the results of regression analysis for a seven variable model, including all second-order cross terms and the third-order three-way cross term (tab. 4), the correlation coefficient is good and the value for the intercept is small. We decide which variables to retain by examining the t-values and their respective probabilities. The probabilities are less than 0.05 for bilirubin, DOPA, the bilirubin-DOPA interaction, and the bilirubin-haemoglobin interaction; thus we retain these variables in the model. The probabilities are greater than 0.05 for haemoglobin, the DOPA-haemoglobin and bilirubin-DOPA-haemoglobin interactions; thus we discard these variables fromthe model. The critical value for F in the partial F test is 4.01, and retention or exclusion of variables based on this test agrees with results of the t-test.

Once one has decided which variables to retain or discard, one must again perform the regression excluding the rejected variables. The interference includes so many different combinations of variables that they all could not be tested at once. Instead, we excluded those with the lowest partial F values and added other combinations of variables that initially appeared less likely to contribute to the regression. After examining several permutations of the variable combinations, we arrived at one set of elements that appeared to include a minimum of variables and yet minimized the mean square of the residuals (tab. 5). This set of variables and combinations includes bilirubin, DOPA, the square of DOPA, and the bilirubin-DOPA, bilirubin-haemoglobin, and bilirubin-DOPA-haemoglobin interactions. We consider all of these coefficients to be significant, on the basis of the values for the t-test and the partial F test (4.01 again being the critical value for F). Another test is whether the difference between this set of variables and combinations (tab. 5) is significantly smaller than the first set (tab. 4). We compared the mean square of the residuals for each set, giving an F ratio of 44.637 : 26.829, which is 1.66. Given the number of degrees of freedom for each regression, the critical value for

Tab. 3. Results of multiple regression analysis of γ-glutamyltransferase-conjugated bilirubin-unconjugated bilirubin interference data.

Correlation coefficient	0.998
Mean square residuals	1.111
F-test[a]	390
Intercept	100.3

Beta coefficient table

Variable	Coefficient	Standard error	t-value	Probability	Partial F
Bc	−1.0	0.068	14.7	0.0001	216
Bu	−0.5	0.068	7.3	0.0007	54
Bc-Bu	−0.025	0.005	4.7	0.0051	22.5

[a] The F-test is based on the ratio of the regression mean square to the residual mean square.

Tab. 4. Results of multiple regression analysis of bilirubin-DOPA-haemoglobin interference data.

Correlation coefficient	0.997
Mean square residuals	44.637
F-test[a]	1231.3
Degrees freedom of residuals[b]	55
Intercept	5.15

Beta coefficient table

Variable[c]	Coefficient	Standard error	t-value	Probability	Partial F
Bilirubin	1.01	0.021	47.8	0.0001	2280
DOPA	36.6	7.20	5.1	0.0001	25.9
Haemoglobin	4.65	3.56	1.31	0.1975	1.70
Bili-DOPA	−0.662	0.042	15.6	0.0001	245
Bili-Hb	−0.082	0.021	3.89	0.0003	15.2
DOPA-Hb	−5.87	7.18	0.818	0.4169	0.669
Bili-DOPA-Hb	0.068	0.042	1.61	0.1133	2.59

[a] The F-test is based on the ratio of the regression mean square to the residual mean square.
[b] One outlier was removed from the data set.
[c] Bili-DOPA stands for the bilirubin-DOPA interaction term, Bili-Hb stands for the bilirubin-haemoglobin interaction term, DOPA-Hb stands for the DOPA-haemoglobin interaction term, and Bili-DOPA-Hb stands for the bilirubin-DOPA-haemoglobin interaction term.

F is 1.55; thus, the value for the F ratio is greater than the critical value, and the set of variables and combinations as shown in table 5 represents a significant improvement over the set shown in table 4.

A function that describes the effect of these interferences on the results is

$$f([Bili], [DOPA], [Hb]) = 4.56 + 0.99[Bili] + 75.3[DOPA] - 0.64[Bili][DOPA] - 0.06[Bili][Hb] - 48.8[DOPA]^2 + 0.04[Bili][DOPA][Hb].$$

One has trouble visualizing such a complicated function. *Contour* plots are easy to draw and interpret. In figure 5 we have plotted this function vs either DOPA or haemoglobin for several different concentrations of bilirubin (the first and third plots). The first plot shows how the degree of curvature of the line depends on the concentrations of bilirubin and DOPA. The slope of these curves is not the same from one concentration of bilirubin to the next. For the haemoglobin interference, the lines do not curve as much as they did for DOPA, but still the slope continues to change. *Contour* plots (the second and third plots of fig. 5) provide an additional perspective. The curves shown in the *Contour* plots represent the bilirubin concentration necessary to maintain a constant result from the instrument as a function of the interferent

Tab. 5. Results of multiple regression analysis of bilirubin-DOPA-haemoglobin interference data based on exclusion of non-significant terms and inclusion of squared terms.

Correlation coefficient	0.998
Mean square residuals	26.829
F-test[a]	2395.8
Degrees freedom of residuals[b]	56
Intercept	4.56

Beta coefficient table

Variable[c]	Coefficient	Standard error	t-value	Probability	Partial F
Bilirubin	0.991	0.013	78.6	0.0001	6175
DOPA	75.3	7.76	9.70	0.0001	94.2
Bili-DOPA	− 0.639	0.025	25.2	0.0001	636
Bili-Hb	− 0.06	0.01	6.053	0.0001	36.6
SquareDOPA	−48.8	7.87	6.20	0.0001	38.5
Bili-DOPA-Hb	0.041	0.02	2.031	0.047	4.126

[a] The F-test is based on the ratio of the regression mean square to the residual mean square.
[b] One outlier was removed from the data set.
[c] Bili-DOPA stands for the bilirubin-DOPA interaction term, Bili-Hb stands for the bilirubin-haemoglobin interaction term, SquareDOPA for the DOPA term squared, and Bili-DOPA-Hb stands for the bilirubin-DOPA-haemoglobin interaction term.

concentration. When the overall effect of an interferent is negative, the bilirubin concentration necessary to maintain the same result increases as the interferent concentration does. Thus, the *Contour* plots curve markedly upwards for both the DOPA and haemoglobin interferences, except at low bilirubin concentrations. The *Contour* plots represent the way the interference problem may present clinically: one may know the approximate interferent concentration and thus could estimate the true bilirubin concentration using the *Contour* plot.

The multiple regression analysis presented in table 6 provides information on the mechanisms of interference, as schematically represented in figures 1 to 4. DOPA by itself may react with the reagents and mimic the bilirubin reaction. DOPA may interact with bilirubin itself or with one of the reaction products of bilirubin with the reagents, thereby decreasing the net absorbance. The statistically significant squared DOPA term implies that DOPA may react with itself, thereby becoming unavailable for the reaction that mimics bilirubin. Because the haemoglobin term by itself was not statistically significant, haemoglobin does not cause a significant absorbance nor does it interact directly with the reagents; however, it does interact with bilirubin. The negative coefficient implies that one possible mechanism is the binding of bilirubin to haemoglobin, thereby decreasing its free concentration. The bilirubin-DOPA-haemoglobin interaction is only marginally statistically significant, and it might be explained by haemoglobin affecting the bilirubin-DOPA interaction and negating a fraction of its negativity.

Conclusions

The graphic model provides us at once with both a diagrammatic scheme of the possible reactions and the basic skeleton of the mechanisms of reaction. The use of Graph Theory allows one to easily write down the combinations of elements that can contribute to the mechanisms. Empirical model-building using response surfaces allows one to use multiple regression analysis to construct a polynomial consisting of the variables and beta-coefficients. The *Taylor* expansion clarifies the meaning of the beta-coefficients, demonstrating that the beta-coefficients are identical to the partial derivatives of the response of the system.

We can determine the statistical significance of the coefficients, retaining those that are statistically significant, while discarding those that are not. Discarding a particular coefficient implies that the arm of the graphic model represented by that coefficient does not contribute to the mechanism. Multiple regression analysis permits the development of an overall mechanism for the interference.

In addition to providing information about the mechanism of interference, multiple regression analysis enables the expression of the interference as a simple polynomial of the concentrations of the analyte and interferents, which includes the interactions between these species. From this polymnomial one can calculate the amount of interference from varying concentrations of analyte and interferents, and assess the clinical significance. This should enable the clinical pathologist to decide which results should no longer be reported. Conversely, given the approximate interferent concentrations, the approximate concentration of analyte can be estimated.

References

1. Kroll, M. H., Ruddel, H., Blank, D. W. & Elin, R. J. (1987) A Model for Assessing Interference. Clin. Chem. *33*, 1121–1123.
2. Letellier, G. (1980) How Does Drug Interference Relate to Quality Assurance? In: *Quality Assurance in Health Care: A Critical Appraisal of Clinical Chemistry* (Rand, R., Eilers, R., Lawson, N. & Broughton, A., eds.) pp. 135–168, American Association for Clinical Chemistry, Washington, D. C.
3. NCCLS (1986) Interference Testing in Clinical Chemistry. Proposed Guideline. [NCCLS Document EP7-P, Vol. 6, No. 13]. NCCLS Villanova, Pa.
4. Huang, C. M., Kroll, M. H., Ruddel, M., Washburn, R. G. & Bennett, J. E. (1988) An Enzymatic Method for 5-Fluorocytosine. Clin. Chem. *34*, 59–62.
5. Gellert, W., Küstner, H., Hellwich, M. & Kästner, H. (1977) *The VNR Concise Encyclopedia of Methematics.* New York: Van Nostrand Reinhold Company. pp. 575–578, 688–692.
6. Box, G. E. P. & Draper, N. R. (1987) Empirical Model-Building and Response Surfaces. New York: John Wiley & Sons. pp. 1–28, 34–99, 219–228.
7. Büttner, J. (1991) Unspecificity and Interference in Analytical Systems: Concepts and Theoretical Aspects. Dtsch. Ges. Klinische Chemie Mitteilungen *22*, 3–11.
8. Kaiser, H. (1972) Zur Definition von Selektivität, Spezifität und Empfindlichkeit von Analysenverfahren. Z. Anal. Chem. *260*, 252–260.
9. Pszonicki, L. (1977) The Specificity Characteristic of Analytical Method – I. Non-Specificity Coefficients. Talanta *24*, 613–616.
10. Pszonicki, L. & Lukszo-Bienkowska, A. (1977) The Specificity Characteristic of Analytical Method – II. Estimation and Experimental Verification of the Non-specificity Coefficients. Talanta *24*, 617–623.
11. Fulks, W. (1969) *Advanced Calculus*, 2nd edn., pp. 259–261, John Wiley & Sons, Inc., New York.
12. Hildebrand, F. B. (1977) *Introduction to Numerical Analysis.* Second edition. New York: Dover Publications, Inc. pp. 6–10.
13. Strang, G. (1980) *Linear Algebra and Its Applications*, 2nd edn., pp. 112–121, Harcourt Brace Jovanovich, Publishers, San Diego.

Dr. Martin H. Kroll
Clinical Pathology Department
National Institutes of Health
Building 10, Room 2C-407
Bethesda, Maryland 20892
U. S. A.